

# A Hybrid Adaptive Compressive Sensing Model for Visual Tracking in Wireless Visual Sensor Networks

Salema Fayed<sup>a</sup>, Sherin Youssef<sup>a</sup>, Amr El-Helw<sup>b</sup>, Mohammad Patwary<sup>c</sup>, and Mansour Moniri<sup>c</sup>

<sup>a</sup> Computer Engineering Department, <sup>b</sup> Electronics and Communication Department

College of Engineering and Technology

AAST, Alexandria, Egypt

<sup>c</sup> Faculty of Computing, Engineering and Technology

Staffordshire University, Stoke on Trent, UK

**Abstract**—The employ of Wireless Visual Sensor Networks (WVSNs) has grown enormously in the last few years and have emerged in distinctive applications. WVSNs-based Surveillance applications are one of the important applications that requires high detection reliability and robust tracking, while minimizing the usage of energy to maximize the lifetime of sensor nodes as visual sensor nodes can be left for months without any human interaction. The constraints of WVSNs such as resource constraints due to limited battery power, memory space and communication bandwidth have brought new WVSNs implementation challenges. Hence, the aim of this paper is to investigate the impact of adaptive Compressive Sensing (CS) in designing efficient target detection and tracking techniques, to reduce the size of transmitted data without compromising the tracking performance as well as space and energy constraints. In this paper, a new hybrid adaptive compressive sensing scheme is introduced to dynamically achieve higher compression rates, as different datasets have different sparsity nature that affects the compression. Afterwards, a modified quantized clipped Least Mean square (LMS) adaptive filter is proposed for the tracking model. Experimental results showed that adaptive CS achieved high compression rates reaching 70%, while preserving the detection and tracking accuracy which is measured in terms of mean squared error, peak-signal-to-noise-ratio and tracking trajectory.

**Index Terms**—Adaptive Compressive Sensing, Compressive sensing, LMS, Surveillance applications, Target tracking, WVSN

## I. INTRODUCTION

Wireless Visual Sensor Networks (WVSNs) have gained significant importance in the last few years and have emerged in several distinctive applications [1],[2]. Due to the evolution of new technologies and techniques, there are immediate needs for automated energy-efficient surveillance systems. WVSN has targeted various surveillance applications in commercial, law enforcement and military purpose as well as traffic control, security in shopping malls and amusement parks. Systems have been developed for video surveillance including highway, subway and tunnel monitoring, in addition to remote surveillance of human activities such as elderly or patients care.

Visual sensor nodes are resource constraint devices bringing the special characteristics of WVSNs such as energy, storage and bandwidth constraints which introduced new challenges [3]. In WVSN large data sets such as video, and still images are to be retrieved from the environment requiring high storage and high bandwidth for transmission. Higher complexity of data processing and analysis is also challenging which are

all quite costly in terms of energy consumption. Furthermore, wireless channels in surveillance applications are subject to noisy conditions; therefore, detection and tracking reliability within such resource constrained condition is the main challenge when designing WVSN surveillance applications. Energy efficient processing and efficient compression techniques are the strongest candidates to overcome such constraints while transmitting data for WVSN applications and hence minimize energy expenditure [2],[4]. Recently, it is very challenging in designing a wireless sensor networks with increased life time [5]. Where a node called WSNMSP430 is developed based on the analysis of the various low power components available in the market and also an energy model was created for processors, transceivers and sensors for predicting the life time of the WSN node.

Much work is present in the literature for surveillance applications within WVSNs [6], [7],[8]. Moreover, there is significant literature for target tracking surveillance applications in WVSN. Kalman filtering [9],[10] is relatively the best linear estimator for target tracking. Kalman filters are robust under optimal conditions, otherwise adaptive approaches are needed to solve these problems which can be either computationally expensive or not always be applicable in real time tracking. To overcome the problems such as changes in the background, occlusion, color, texture and size. A novel combined Gaussian hidden Markov model and Kalman Filter is proposed in [11] for multiple target detection and tracking Surveillance applications

Particle filtering which is known to be suitable for real time tracking and non-linear non-Gaussian processes, it relies on motion parameter estimation and probability estimates [12]. Subsequently, the performance of the particle filter in terms of tracking reliability decreases with noisy or low resolution frames and with false positive detection of target [10]. Classical active contour [13] for target tracking fails in tracking multiple targets at once so occlusion problems are difficult to solve. In [14], the active contour is modified to resolve occlusion problem by performing merging and splitting when two targets get close together or move apart. However, there is a probability that the target is lost if the displacement of the target between two consecutive frames is large. Least Mean Square (LMS) algorithm is relatively

simple, has much lower computational complexity than the original Kalman filters and other adaptive algorithms; it does not require correlation function calculation nor does it require matrix inversions. Moreover, it is suitable for real time image applications [15],[16].

Based on the above literature, to attain a trade off between computational complexity and detection and tracking accuracy in the context of energy constrained WWSN, an image processing scheme is required with optimal pre-processing and post-processing can provide intended target detection and tracking accuracy within energy constraint nature of WWSN. Moreover, high volume data sets acquired in WWSN surveillance applications, should be represented in such a way that it requires optimum storage, energy, and allow reliable transmission due to the constraint on the physical and radio resources. In a surveillance application within WWSN, an image is captured and required to be sampled for storage as well as to be transmitted through wireless channel. According to Shannon-Nyquist sampling theory the minimum number of samples required to accurately reconstruct the signal without losses is twice its maximum frequency [17]. It is always challenging to reduce this sampling rate as much as possible, hence reducing the computation energy and storage. Recently proposed Compressive Sensing (CS) [17] is expected to be a strong candidate to overcome the above mentioned limitations where CS has been considered for different aspects of surveillance applications due to its energy efficient and low power processing as reported in [18],[19].

CS theory shows that a signal can be reconstructed from far fewer samples than required by Nyquist theory as it is always challenging to reduce the sampling rate as possible, provided that the signal is sparse (where most of the signal's energy is concentrated in few non-zero coefficients) or compressible in some basis domain [20].

In [21] a new method of facial expression recognition based on the sparse representation classifier is presented where CS has been proposed as an efficient classification method. The newly-emerged CS theory has been used to form a new classification technique showing promising performance on pattern recognition. In [18], compressive sensing for background subtraction and multi-view ground plane target tracking are proposed. A convex optimization known as basis pursuit or orthogonal matching pursuit is exploited to recover only the target in the difference image using the compressive measurements to eliminate the requirement of any auxiliary image reconstruction. Other work in compressive sensing for surveillance applications has been proposed in [22], where an image is projected on a set of random sensing basis yielding some measurements. In [19] a novel compressive particle filter for tracking one or more targets in video is presented using a reduced set of observations. It is shown that, by applying compressive sensing ideas in a multi-particle-filter framework, it is possible to preserve tracking performance while achieving considerable dimensionality reduction, avoiding costly feature extraction procedures. Additionally, the target locations are predicted directly, without the need to reconstruct each image. However, the proposed algorithm

failed to provide acceptable performance for fast moving targets. In addition, it is not designed for WWSN applications thus constraints of WWSN such as energy and memory constraints were not taken into consideration.

Another promising direction is the adaptive CS, in [23], energy efficient data collection in WSN using adaptive compressive sensing is proposed. An adaptive approach is proposed to select a routing path by choosing sensors required to transmit their data. However, in this approach adaptive CS is only applied for sensor nodes selection and no compression is performed on the transmitted data. A heuristic to solve the optimization problem (which is proven NP-hard) is proposed in [24] to find a measurement matrix that maximizes the information gain per energy expenditure. It was shown that under suitable conditions, one can reconstruct an  $(N \times N)$  matrix of rank  $r$  from a small number of its sampled measurements. This is done by solving an optimization problem, provided that the number of measurements is of order of  $N^{1.2r} \log n$ , exact matrix recovery would be guaranteed with a reduced number of measurements. In [25, 26], an adaptive approach to compressed sensing is proposed using a single pixel camera. Instead of using a representation (such as pseudo-random binary masks) that is incoherent with a conventional transform (as wavelets) to acquire the visual data. The image is sampled directly in the wavelet domain by tuning the Digital Micro-Mirror Device (DMD) of the single pixel camera to directly collect only the significant wavelet coefficients.

Most of the CS algorithms proposed [18, 19, 22, 27] are non-adaptive which means the random measurement matrix is not chosen according to information collected. An important issue is to make the measurement matrix adaptive. Subsequently, most existing work in adaptive compressive sensing use heuristic techniques which are computationally expensive, hence taking only into consideration the accuracy of the approximate data field without considering the energy factor. Therefore, considering the resource constraint within WWSN for surveillance applications, the feasibility of such feature specific adaptation of CS for reliable target detection and tracking is the major focus of the proposed investigations. Hence, in this paper, the impact of adaptive CS is investigated in designing target detection and tracking techniques for WWSNs-based surveillance applications, without compromising the energy constraint which is one of the main characteristics of WWSNs. Adaptive CS is expected to reduce the size of sampled data with low complexity processing due to its low power simple process [20], hence saving space, energy of processing and transmission as well as channel bandwidth. Hence, a compressive sensing-based single/multi target tracking using LMS is proposed which is expected to reduce energy consumption, space requirement and communication overhead, with acceptable tracking reliability which will be represented as minimal mean square error (MSE).

The rest of the paper is organized as follows, Introduction to CS is presented in Section II. Section III presents the proposed system model. The proposed technique for adaptive CS-based target tracking is given in Section IV. Simulations and results

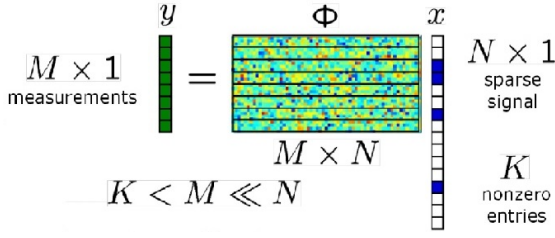


Fig. 1. Compressive sensing measurement process

are provided in Section V and finally the conclusion in Section VI.

## II. COMPRESSIVE SENSING THEORY

Suppose image  $X$  of size  $(N \times N)$  is  $K$ -sparse that either sparse by nature or sparse in  $\Psi$  domain, CS exploits the sparsity nature of frames, so it compresses the image using far fewer measurements [28],[20],[29]. Although, it is not necessary for the signal itself to be sparse but compressible or sparse in some known transform domain  $\Psi$  according to the nature of the image, smooth signals are sparse in the Fourier basis, and piecewise smooth signals are sparse in a wavelet basis.  $\Psi$  is the basis invertible Orthonormal function of size  $(N \times N)$  driven from a transform such as the DCT, fourier, or wavelet, where  $K \ll N$ , that is, only  $K$  coefficients of  $x$  are nonzero and the remaining are zero, thus the  $K$ -sparse image  $X$  is compressible. CS then guarantees acceptable reconstruction and recovery of the image from lower measurements compared to those required by shannon-Nyquist theory as long as the number of measurements satisfies a lower bound depending on how sparse the image is. Hence,  $X$  can be recovered from measurements of size  $M$  where  $M \geq K \log N \ll N$ . Eq.(1) shows the mathematical representation of  $X$

$$X = \Psi S \quad (1)$$

$S$  contains the sparse coefficients of  $X$  of size  $(N \times N)$ ,  $s_i = \langle X, \psi_i^T \rangle = \psi_i^T X$ ,  $S = \Psi^T X$ . The image is represented with fewer samples from  $X$  instead of all pixels by computing the inner product between  $X$  and  $\Phi$ , namely through incoherent measurements  $Y$  in Eq.(2), where  $\Phi$  is a random measurement matrix of size  $(M \times N)$  where  $K \ll M \ll N$ . Fig.1 shows the CS measurement process [30].

$$y_1 = \langle x, \phi_1 \rangle, y_2 = \langle x, \phi_2 \rangle, \dots, y_m = \langle x, \phi_m \rangle.$$

$$Y = \Phi X = \Phi \Psi S = \Theta S \quad (2)$$

Since  $M < N$ , recovery of the image  $X$  from the measurements  $Y$  is undetermined, However, if  $S$  is  $K$ -sparse, and  $M \geq K \log N$  it has been shown in [20] that  $X$  can be reconstructed by  $\ell_1$  norm minimization with high probability through the use of special convex optimization techniques without having any knowledge about the number of nonzero coefficients of  $X$ , their locations, neither their amplitudes which are assumed to be completely unknown a priori [29],[28],[31]

$$\min \|\hat{X}\|_{\ell_1} \text{ subject to } \Phi \hat{X} = Y \quad (3)$$

Convex optimization problem can be reduced to linear programming known as Orthogonal Matching Pursuit (OMP) which was proposed in [32] to handle the signal recovery problem. It is an attractive alternative to Basis Pursuit (BP) [33] for signal recovery problems. The major advantages of this algorithm are its speed and its ease of implementation. As seen, the CS is a very simple process as it enables simple computations at the encoder side (sensor nodes) and all the complex computations for recovery of frames are left at the decoder side or BS.

## III. SYSTEM MODEL

This work proposes an adaptive compressive sensing model which is expected to reduce space requirements and communication overhead with low processing complexity while preserving detection and tracking accuracy.

Consider for a surveillance application a WVSAN model composed of  $V$  visual sensor nodes and one or more BS. Each sensor node  $i$  is required to capture images from a video sequence and detect the presence of objects. At the time where a sensor node enters a 'wake-up' state, the time reference for the frame count is assumed to be  $t = 0$ . Hence, a single snapshot at  $t = 0$  is expected to be stored within the memory allocated at the sensor node; that is assumed to be the background for the intended target tracking; denoted as  $X_b$ . The following frames are the subsequent captured frames  $X_t$  with  $t > 0$ . Hence,  $X_b$  and  $X_t$  are the background and test images respectively of size  $(N \times N)$  each. Let us assume most features of the targets are known to the monitoring center. However, the existence and the location of targets are required for monitoring. The receiver or BS also has prior explicit information of the background. To achieve higher compression rates, the foreground target is extracted first by background subtraction resulting in the difference frame. Hence, assuring sparsity as the difference frame is always sparse regardless the sparsity nature of real frames. Within the image frame, The extraction of foreground target  $X_d$  is achieved at each sensor node where adaptive CS is then applied for transmission through the wireless channel. CS adaptively chooses the compression rate according to the sparsity nature of difference frames which varies from one dataset to another. The training/calibration phase is pre the CS phase and is discussed later in Sec.IV-B. At the BS side, the receiver decompresses the received compressed data obtaining  $\hat{X}_t$  to predicts the intended target's next location for tracking. The system model for the proposed WVSAN is shown in Fig.

2

## IV. PROPOSED ADAPTIVE CS-BASED TRACKING ALGORITHM

### A. Foreground detection and morphology operations

At each sensor node, after each image frame is being captured, some preprocessing might be required. In our case, to assure sparsity within the image frame, the foreground target is extracted first based on thresholding the absolute difference between current frame  $X_t$  and background frame  $X_b$ ,  $X_d = |X_t - X_b| > \gamma$ , where  $\gamma$  is a given threshold to

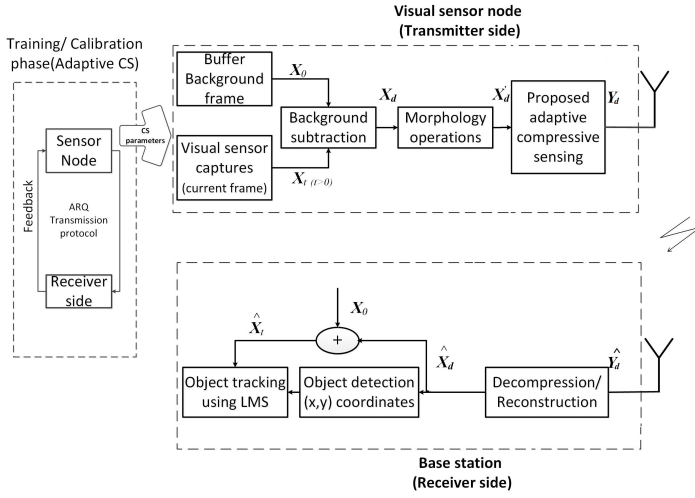


Fig. 2. The proposed model for WVSN-based surveillance application

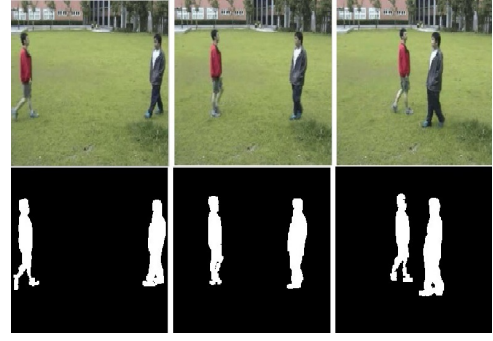
extract the foreground target by background subtraction resulting in the difference frame  $X_d$ . Hence, instead of producing the compressed measurements for  $X_b$  and  $X_t$  separately, the compressed measurements are produced directly for  $X_d$ , as the difference frame is always sparse regardless of the sparsity nature of real frames.

Once the foreground is detected, morphology operations [34] such as erosion and dilation operations are then applied for noise removal and blob formation respectively. The purpose of morphological processing is primarily to remove imperfections added during segmentation, in the context of our work, after background subtraction an opening or closing operations are then applied depending on the nature of images. Sometimes an opening operation is performed where erosion is first applied as a noise removal method by applying the specified structuring element to remove unwanted pixels, followed by dilation to fill the holes within target objects forming a connected object blob by linking the unconnected parts of the target. Hence, any regions that have survived the erosion are restored to their original size by the dilation. Or a closing operation, obtained by dilation of the image using the specified structuring element to form a connected object blob, followed by erosion of the resulting image to restore the original size of objects. It can fill holes in the regions while keeping the initial region sizes[34]. Fig.3 shows the blob formation after background subtraction and morphological operations

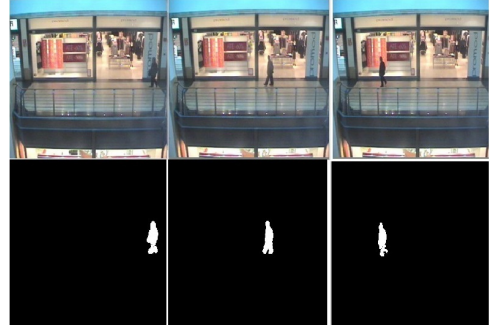
### B. Proposed adaptive Compressive Sensing

After the foreground blob  $X_d$  is being extracted from the difference image, the proposed adaptive CS is then applied to  $X_d$  by multiplying it by a random projection sensing matrix  $\Phi$  producing the compressed measurements  $Y_d$ . At the BS side, the received compressed data is decompressed for the reconstruction of the estimated data  $\hat{X}_d$ . As mentioned,  $X_b$  is known to the BS, making it possible to reconstruct the original test frame  $\hat{X}_t$  by adding  $X_b$  to  $\hat{X}_d$ .

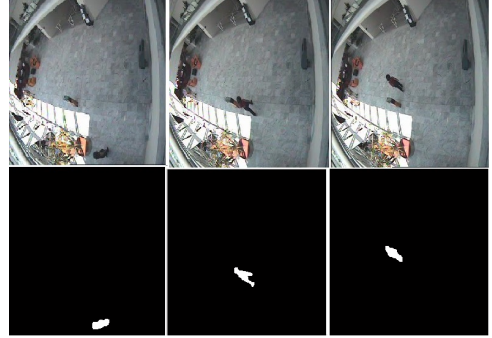
For any given scheme, different  $M$  and  $\Phi$  are needed, as



(a) Walking men



(b) Shopping center 1



(c) Shopping center 2

Fig. 3. First row in (a),(b) and (c) shows test frames and background subtraction results and blob formation in second row

stated earlier the value of  $M$  is inversely proportional to the degree of sparsity of an image. If the same value of  $M$  is used for all different schemes, it is expected that the reliability of target detection will be different as the degree of sparsity varies from one image to another. For this reason there is a great challenge for adaptive CS by making  $M$  variable depending on how sparse the image is. For the adaptive CS, the CS process is preceded by a calibration phase. During that phase an Automatic Repeat Query (ARQ) transmission protocol is used between sensor nodes and the receiver side, as a feedback is needed for the adaptation phase. Initially, an arbitrary value of  $M$  is chosen according to a sparsity



measure and is used to obtain the compressed measurements  $Y_d$ . The sensor node is then set to transmit  $Y_d$  to the receiver side where the image is to be reconstructed, and based on the reconstruction error a decision is made whether the reconstruction is satisfactory or not. In case the reconstruction results are satisfactory, the receiver node sends a 'zero' flag through the feedback channel ending the calibration phase; otherwise a 'one' flag is to be sent. While the sensor node receives a 'one' flag, it is expected to change the value of  $M$  and change  $\Phi$  accordingly, the sensor node repeats the search for an optimum value of  $M$  at the CS adaptation process till it receives a zero feedback from the receiver. At this point, the optimum values for  $M$  and  $\Phi$  obtained are used next in the CS process. Fig.4 shows a flow chart summarizing the entire adaptive CS process. Below are the steps undertaken during the entire process

- **Step 1:**  $X_d = |X_t - X_b| > \gamma$ , where  $\gamma$  is a given threshold to extract the foreground target
- **Step 2:**  $\Phi$  is a randomly chosen sensing matrix from the adaptive process of size  $M \times N$ , where  $M \ll N$
- **Step 3:** produce the compressed measurements  $Y_d = \Phi X_d$
- **Step 4:** sensor nodes transmits  $Y_d$  through the wireless channel
- **Step 5:** at the receiver side,  $\Phi$  must be known for the decompression of  $Y_d$ .  $\hat{X}_d$  is reconstructed from the compressed measurements  $Y_d$ , resulting in a frame with only the foreground target present.
- **Step 6:** the real frame  $\hat{X}_t$  is then obtained by adding  $\hat{X}_d$  to the background frame  $X_b$  which is also has to be known to the receiver side apriori.
- **Step 7:** the targets locations are obtained after reconstructing the real frame producing a trajectory for the complete path of each moving target

### C. Least Mean Square (LMS) tracking

The LMS algorithm, is referred to as adaptive filtering algorithm since the statistics are predicted continuously, hence it can adapt to changes. LMS incorporates an iterative procedure during the training phase where it predicts the required coefficients to minimize the mean square error (MSE). This is accomplished through successive corrections to the expected set of coefficients which eventually leads to the minimum MSE.

The outputs are linearly combined after being scaled using corresponding weights. The weights are computed using LMS algorithm based on MSE criterion. Therefore the spatial filtering problem involves estimation of a signal from the received signal, by minimizing the error between the reference signal, which closely matches or has some extent of correlation with the desired signal estimate and the output. The LMS algorithm is initiated with an arbitrary value  $w(0)$  for the weight vector at  $n = 0$ . The successive corrections of the weight vector eventually leads to the minimum value of the mean squared error. The weight update can be given by the following equation

$$w(n+1) = w(n) + \mu x(n)e(n) \quad (4)$$

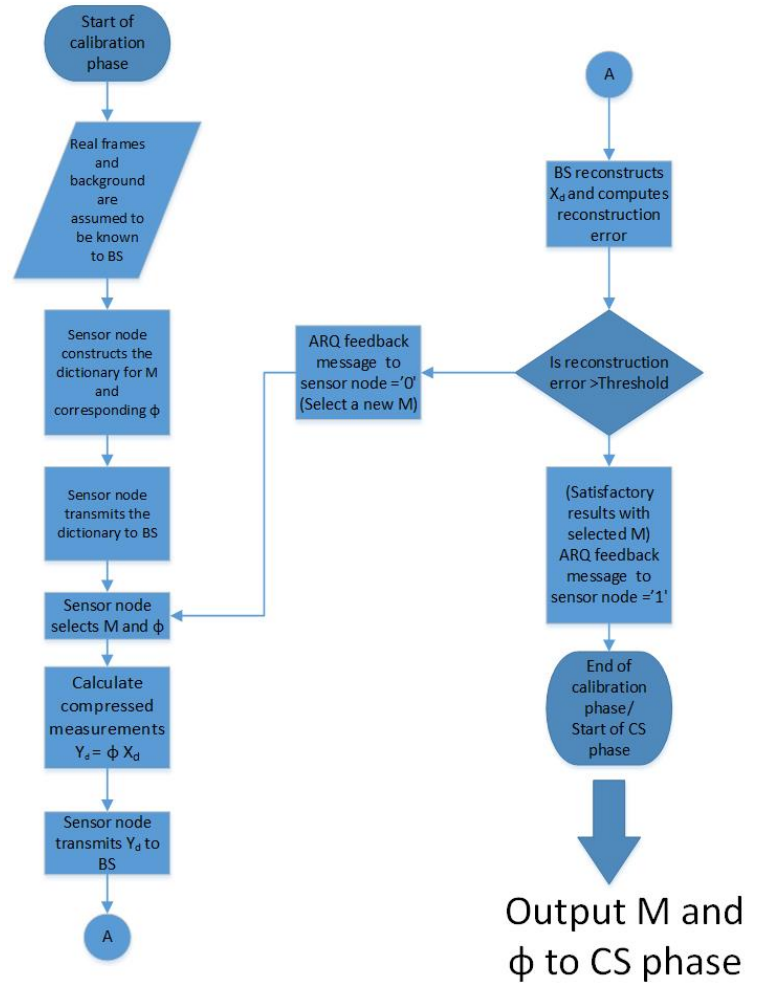


Fig. 4. Flowchart for the adaptive CS process

where,  $x(n)$  is the input signal,  $\mu$  is the step size parameter,  $e(n)$  is the MSE between the predicted output  $y(n)$  and the reference signal  $d(n)$  which is given by

$$e(n) = (d(n) - y(n))^2 \quad (5)$$

the output  $y(n)$  is calculated as follows

$$y(n) = x(n)w(n) \quad (6)$$

$\mu$  is selected by the autocorrelation matrix of the filter inputs. In other words, the tap-weights can converge to an optimum result if and only if the step-size parameter  $\mu$  is selected as  $0 < \mu < 1/\lambda_{max}$

where,  $\lambda_{max}$  is the maximum eigenvalue of the autocorrelation matrix which has a relationship of the input signal  $x(n)$ . The smallest the eigen value spread the faster the convergence rate. Eigen value spread is defined as the ratio between the maximum and minimum eigen values. The LMS algorithm simplifies the estimation of autocorrelation matrices by using the instantaneous values of the autocorrelation matrices instead of their actual values.

There are several variants of the LMS algorithm present in the literature [35–37] to deal with the shortcoming of its

basic form and aim for lower computational complexity and faster adaptation processes. For the proposed model, target tracking is achieved using a modified quantized clipped LMS technique to predict the target's next location, by modifying the "sgn" function as shown below with predefined threshold values;  $D_1$  and  $D_2$  used to clip the input data.

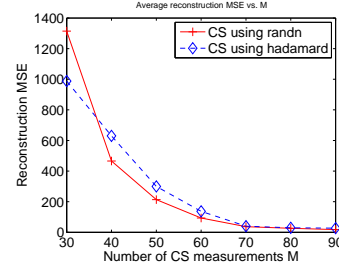
$$mqsgn(\mathbf{x}(n)) = \begin{cases} 1 & \mathbf{x}(n) > D_1 \\ 0 & -D_2 < \mathbf{x}(n) < D_1 \\ -1 & \mathbf{x}(n) < -D_2 \end{cases}$$

## V. SIMULATIONS AND RESULTS

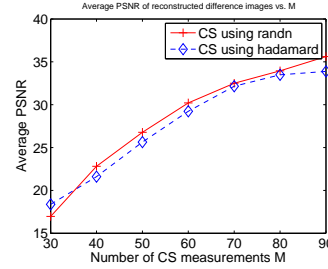
Based on the system model proposed, simulations and experiments are conducted to evaluate the performance of the adaptive CS-based target detection and tracking algorithm. Simulations are performed for the WVSN-based surveillance application in both outdoor and indoor scenes for single and multi-target tracking. Background and target's appearance are assumed to be static to investigate the effect of adaptive CS on the detection and tracking algorithms, hence schemes are chosen to reflect this assumption. Moreover, to illustrate the relation between the number of measurements required for adaptive CS to guarantee reconstruction and how sparse the image is. Simulations are performed on different schemes with different sparsity levels as in Fig.3; for the "outdoor scheme", "Walking men" is chosen to resemble multi target tracking captured by [38]. While "indoor scheme", "Shopping center 1" and "Shopping center 2" filmed for the EC funded CAVIAR project found in [39] for different indoor scenes tracking a single target.

Mean square error (MSE) and peak signal to noise ratio (PSNR) are used as performance indicators to test the reliability of adaptive CS. MSE and PSNR are compared for different number of CS measurements  $M$ , where the MSE is the reconstruction error measured between real and reconstructed frames and PSNR is measured after frames recovery to reflect the quality of image reconstruction which will later on reflects the ability of reliable tracking. The background frame and  $\Phi$  are known to the receiver node. Two candidate sensing matrices have been compared; normally distributed random numbers using Matlab function "randn" and a walsh-hadamard. Although the measurements are defined by a matrix multiplication, the operation of matrix-by-vector multiplication is seldom used in practice, because it has a complexity of  $O(MN)$  which may be too expensive for real time applications. When a randomly permuted Walsh-Hadamard matrix is used as the sensing matrix, the measurements may be computed by using a fast transform which has complexity of  $O(K \log(N))$  [40]. The Hadamard matrix, is an  $(N \times N)$  square matrix whose entries are either +1 or -1 and whose rows are mutually orthogonal, the matrix is first randomly reordered then,  $M$  samples are randomly chosen to construct the  $(M \times N)$  random sensing matrix  $\Phi$ .

As stated earlier, the ability of reliable tracking depends on acceptable recovery of images. In other words, if CS fails in image reconstruction the targets location can not be detected. Hence, choosing the right value of  $M$  is critical in image



(a) Reconstruction MSE

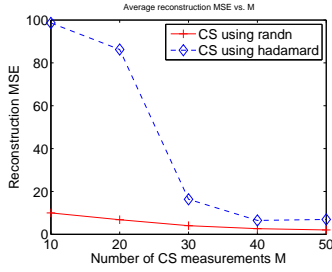


(b) PSNR

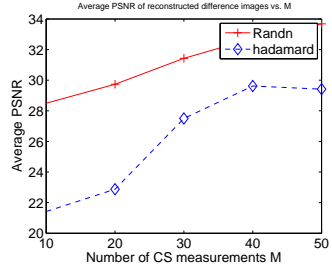
Fig. 5. Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Walking men outdoor scheme"

reconstruction and afterwards tracking. It is clear from the results in Fig.5 for the outdoor scheme and Fig.6 and 7 for the indoor schemes that for different sparsity levels different values of  $M$  and compression rates are required. When reaching optimum value of  $M$  least MSE while preserving a 33dB PSNR. For illustration, MSE decreases as  $M$  increases till reaching the optimum value, it has been shown that the lower bound on  $M$  is depending on how sparse the difference frame  $X_d$  is or in other words proportional to the ratio between the number of non-zero coefficients and the total number of pixels in a frame. For "outdoor scheme", adaptive CS sets  $M$  to 90 in Fig.5(a) to achieve satisfactory results. While for "indoor scheme", it is obvious from Fig.6(a) and 7(a) that for single-target tracking (where there is lower number of non-zero coefficients), better MSE is achieved with lower  $M$  for the "indoor scheme", reduced to 50 and 60 for "Shopping center 1" and "Shopping center 2" respectively, compared to multi-target tracking while maintaining least MSE and 33dB PSNR as in Fig.6.

As for MSE, Fig.5(b), 6(b) and 7(b) show the effect of  $M$  on PSNR for the different schemes. For each scheme, according to the sparsity nature of each scheme, the number of measurements  $M$  required will differ to obtain guaranteed reconstruction which is defined here in terms of PSNR. For low values of  $M$  it is hard to achieve a good PSNR, to reach the acceptable value,  $M$  should increase till reaching its optimum value as discussed earlier. To illustrate this for the "indoor scheme", to achieve a PSNR of  $\approx 33dB$ ,  $M$  reached  $\approx 55$ , while for the "outdoor scheme" if the same  $M$  is used, we could not attain a PSNR higher than 25dB.



(a) Reconstruction MSE

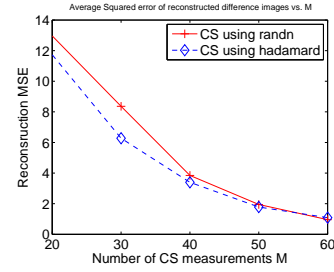


(b) PSNR

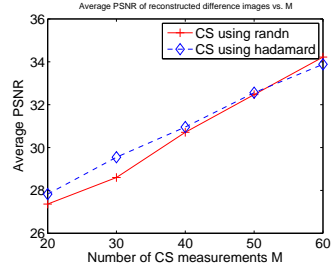
Fig. 6. Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Shopping center 1 indoor scheme"

The above simulation were carried out using two different sensing matrices, Randn and walsh-Hadamard. They are compared with respect to MSE and PSNR as in Fig.5, 6 and 7. It is clear from the results that when reaching the optimum value of  $M$  both sensing matrices perform nearly the same except in some cases in Fig.6 shows that Randn gives slightly a better performance than Hadamard. But this can be negligible when compared to the reduction in complexity gained by using Hadamard matrix which helps in accomplishing the main objective to save sensor nodes power and as a result maximizes their lifetime.

Fig.8 and 9 summarize and demonstrate the effect of the target size ratio on the number of measurements  $M$  needed in terms of reconstruction MSE and PSNR (the target size ratio is expressed as a ratio between non-zero pixels representing the target and the total size of the image frame, which reveals how much space the target acquires and how sparse the image is). It is clear from Fig.8 that for smaller target sizes, lower values of  $M$  are used while at the same time achieving the least MSE and PSNR of  $\approx 33dB$  as in Fig.9(a) and 9(b), respectively. While for larger target sizes, a higher  $M$  is required to achieve the same performance achieved for frames with smaller targets. Experiments were carried out using the same  $M$  set to 50 for the different schemes (different sparsity levels). For example, frames with small size targets gave better reconstruction results in terms of least MSE and a  $33dB$  PSNR as in Fig.9(a) and 9(b). Whereas, if the targets size grew bigger such as acquiring 60% space of the total frame size, with  $M$  set constant reconstruction results in high MSE and only  $18dB$  PSNR. In that case  $M$  should be set to 90 or



(a) Reconstruction MSE



(b) PSNR

Fig. 7. Comparing reconstruction MSE and PSNR using randn and walsh sensing matrices for "Shopping center 2 indoor scheme"

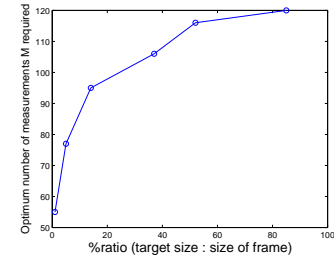
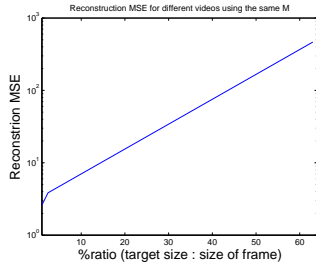


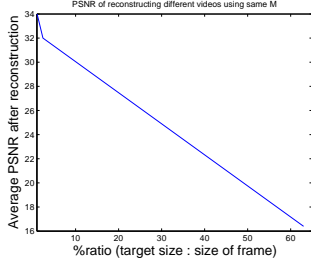
Fig. 8. Relation between the percentage ratio of target size:frame size vs.  $M$

higher based on the sparsity nature to reach a low MSE and a PSNR of  $\approx 30dB$  that was attained by lower  $M$  ( $M = 50$ ) when compressing frames with targets of size  $< 10\%$  of the frame size. These results reflect the constraint of the lower bound of  $M$  discussed in sec.II and give a key to the problem when  $M$  is required to be kept as small as possible. Where in that case the size of targets is controlled by zooming or changing the location of sensor nodes while bearing in mind to keep the scene of interest in the camera's field of view. By taking snapshots from a further location the total space acquired by the target is hence reduced and as a result  $M$  can be reduced, and the goal of reducing the size of transmitted data is met .

Another performance indicator is the correlation coefficient. After reconstructing the compressed measurements, the correlation coefficient indicates how likely the reconstructed frame correlates with the original one. Fig.10 shows by increasing  $M$  till reaching its optimum values the correlation coefficients



(a)



(b)

Fig. 9. Relation between the percentage ratio of target size:frame size and (a) reconstruction MSE, (b) average PSNR

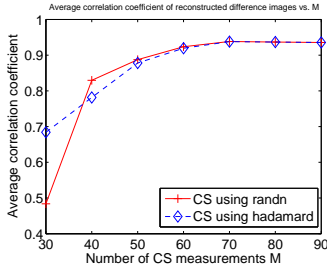


Fig. 10. Correlation coefficient for different M

is nearly 100%, this implies that adaptive CS has not affected the image quality after recovery, whereas less number of measurements were required reducing the size of transmitted data.

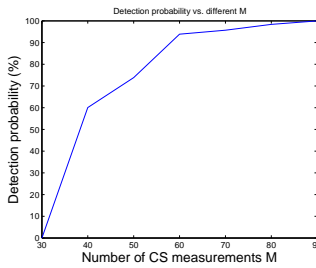
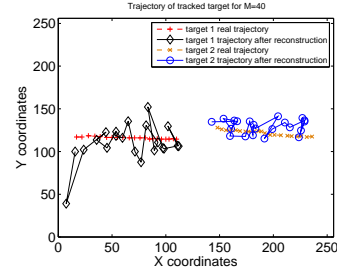


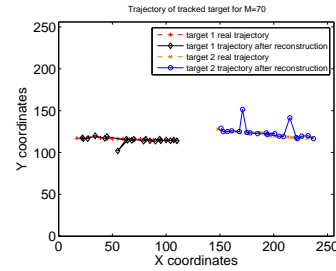
Fig. 11. Probability of detection vs different values of M

Fig.11 shows the probability of detection for different values of measurements  $M$ , it is clear from the graph that for lower values of  $M$  the target is misdetected. This reflects

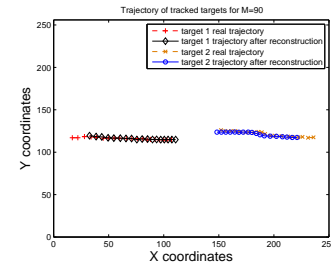
the fact that the reconstruction can not be guaranteed with lower values of  $M$ . The probability of detection increases till reaching 100% as  $M$  increases to its optimum value selected during the adaptive CS process.



(a) M=40



(b) M=70

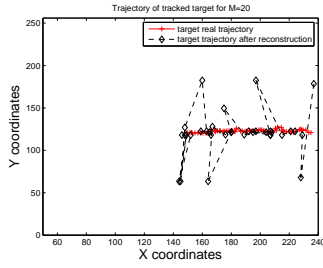


(c) M=90

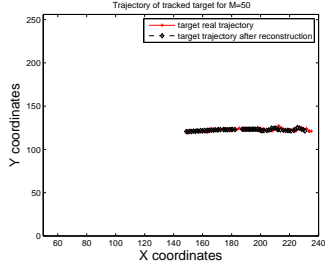
Fig. 12. Comparing predicted trajectory of multi-targets using LMS for "Walking men" (using different M for CS)

CS states that when enough measurements are used for compression, the reconstruction is done with high accuracy depending on a lower bound of  $M$ . Trajectory tracking of moving targets is considered to reflect the degree of reconstruction accuracy. Tracking reliability is tested by comparing the moving target's real and predicted trajectories using LMS. Fig.12, 13 and 14 show the  $(x,y)$  position plots of the path tracked for the targets in the camera's scene. Fig.12(a) and 12(b) show that (for "Walking men") for lower values of  $M <$  optimum value (40 and 70 respectively), frames can not be reconstructed properly and as a result the targets tracks are not matching their real trajectories, whereas for optimum values of  $M$  reaching 90, LMS accurately predicted the target's locations and the results are closely matching the real target





(a) M=20



(b) M=50

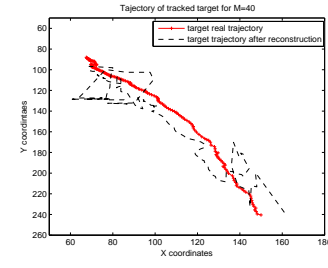
Fig. 13. Comparing predicted trajectory of single target using LMS for "Shopping center 1" (using different M for CS)

trajectory before compression. Fig.13 and 14 illustrate the same for the "Shopping center 1" and "Shopping center 2", respectively.

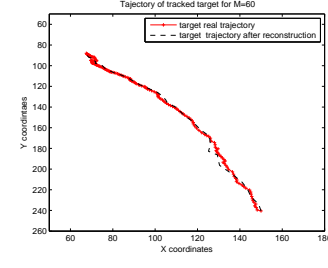
## VI. CONCLUSION

In this paper, the constraints of WVSNs are characterized such as resource constraints due to limited battery power, memory space and communication bandwidth. These constraints brought new implementation challenges to investigate adaptive CS in designing robust target detection and tracking techniques for surveillance applications without compromising the tracking performance as well as space and energy constraint. CS has been expected to be a strong candidate to achieve high compression rate using simple computations. Since the compression rates differ from one dataset to another depending on the degree of sparsity. An adaptive CS technique has been proposed and has proven to achieve high compression rates with minimum reconstruction error.

Experiments were carried out to evaluate the performance of adaptive CS and its effect on target detection and tracking. Simulations have shown that CS is a strong candidate to reduce the size of images without degrading the tracking performance. Results have shown that using adaptive CS up to 31% measurements of data are required to be transmitted, while preserving the reconstruction quality which is measured in terms of MSE, PSNR and trajectory tracking. The reconstruction MSE adaptively decreases till reaching the lower bound on the number of compressed measurements while preserving the acceptable PSNR. In addition, for different



(a) M=40



(b) M=60

Fig. 14. Comparing predicted trajectory of single target using LMS for "Shopping center 2" (using different M for CS)

schemes where the sparsity nature of each image differs, adaptive CS chooses the compression rates accordingly. Moreover, surveillance application within WVSNs is one of the important applications that requires high detection reliability and robust tracking. After image reconstruction, the impact of adaptive CS on target tracking is investigated using LMS adaptive filter to predict target's next location. Target's trajectory tracking has been used as a performance indicator for the LMS algorithm. Results have demonstrated that the predicted path closely matches the target's real path which illustrates the accuracy of LMS and that adaptive CS has not affected the performance of target detection and tracking.

## REFERENCES

- [1] A.Sharif, V.Potdar, and E.Chang, "Wireless multimedia sensor network technology: A survey," in *Proceedings of Industrial Informatics, 7th IEEE International Conference, 2009. INDIN 2009.*, June 2009, pp. 606–613.
- [2] I.F.Akyildiz, T.Melodia, and K.R.Chowdhury, "A survey on wireless multimedia sensor networks," *computer Networks*, vol. 51, pp. 921–960, March 2007.
- [3] S.Soro and W.Heinzelman, "A survey on visual sensor networks," *Hindawi publishing corporation, Advances in Multimedia*, vol. 2009, no. 640386, pp. 1–21, May 2009.
- [4] Y.Charfi, B.Canada, N.Wakamiya, and M.Murata, "Challenges issues in visual sensor networks," in *IEEE on wireless Communications*, April 2009, pp. 44–49.
- [5] G. Kannan and T. R. Raja, "An efficient cluster-based reliable power aware scheme (rpas) for network longevity

- in wsn,” *Wseas Transactions on Computers*, vol. 12, no. 9, pp. 366–373, September 2013.
- [6] X.Wang, S.Wang, and D.Bi, “Distributed visual-target-surveillance system in wireless sensor networks,” *IEEE Transactions on Systems, MAN, and Cybernetics*, vol. 39, no. 5, pp. 1134–1146, October 2009.
  - [7] X.Wang, S.Wang, D.W.Bi, and J.J.Ma, “Distributed peer-to-peer target tracking in wireless sensor networks,” *MDPI, open access journal on the science and technology of sensors and biosensors*, vol. 7, pp. 1001–1027, 2007.
  - [8] X.Wang and S.Wang, “Collaborative signal processing for target tracking in distributed wireless sensor networks,” *Elsevier journal on Parallel and distributed computing*, vol. 67, p. 501–515, 2007.
  - [9] D. Simon, “Kalman filtering with state constraints: a survey of linear and nonlinear algorithms,” *The Institution of Engineering and Technology, Control theory applications*, vol. 4, no. 8, pp. 1303–1318, 2010.
  - [10] J.C.Noyer, P.Lanvin, and M.Benjelloun, “Non-linear matched filtering for object detection and tracking,” *Elsevier Pattern Recognition Letters*, vol. 25, pp. 655–668, 2004.
  - [11] S.Vasuhi and V.Vaidehi, “Target detection and tracking for video surveillance,” *Wseas Transactions on Signal Processing*, vol. 10, pp. 168–177, 2014.
  - [12] A.C.Sankaranarayanan, A.Veeraraghavan, and R.Chellappa, “Object detection, tracking and recognition for multiple smart cameras,” in *Proceedings of the IEEE*, vol. 96, no. 10, October 2008, pp. 1606–1624.
  - [13] S. Lefevre and N. Vincent, “Real time multiple object tracking based on active contours,” September 2004.
  - [14] J.Malcolm, Y.Rathi, and A.Tannenbaum, “Multi-object tracking through clutter using graph cuts,” in *The International Conference on Computer Vision (ICCV)*, 2007.
  - [15] S.Haykin, *Adaptive Filter Theory*. Prentice Hall, 2002, vol. 0-13-048434-2, ch. Least mean square adaptive filters, pp. 231–247.
  - [16] P.S.R.Diniz, *Adaptive Filtering*. The Springer International Series in Engineering and Computer Science, January 1997, vol. 399, ch. The Least-Mean-Square (LMS) Algorithm, pp. 79–135.
  - [17] R. Baraniuk, “Compressive sensing,” *IEEE Signal Processing Magazine*, pp. 118–124, July 2007.
  - [18] V.Cevher, A.Sankaranarayanan, M. Duarte, D.Reddy, R. Baraniuk, and R.Chellappa, “Compressive sensing for background subtraction,” 2008.
  - [19] E. Wang, J. Silva, and L. Carin, “Compressive particle filtering for target tracking,” in *IEEE/SP 15th Workshop on Statistical Signal Processing, SSP*, September 2009, pp. 233–236.
  - [20] J.Romberg, “Imaging via compressive sampling,” *IEEE Signal Processing Magazine*, pp. 14–20, March 2008.
  - [21] S. Zhang, X. Zhao, and B. Lei, “Facial expression recognition using sparse representation,” *WSEAS Transactions on Systems*, vol. 11, no. 8, pp. 440–452, August 2012.
  - [22] A.Mahalanobis and R.Muise, “Object specific image reconstruction using a compressive sensing architecture for application in surveillance systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1167–1180, July 2009.
  - [23] C.T.Chou, R.Rana, and W.Hu, “Energy efficient information collection in wireless sensor networks using adaptive compressive sensing,” in *IEEE 34th Conference on Local Computer Networks (LCN)*, Zurich, Switzerland, October 2009, pp. 443–450.
  - [24] E.J.Candes and B.Recht, “Exact matrix completion via convex optimization,” *CoRR*, vol. abs/0805.4471, 2008.
  - [25] S. Deutsch, A. Averbuch, and S. Dekel, “Adaptive compressed image sensing based on wavelet modeling and direct sampling,” in *Proceedings of the 8th International Conference on Sampling Theory and Applications, Marseille, France*, 2009.
  - [26] S. Dekel, “Adaptive compressed image sensing based on wavelet-trees,” online. Available: <http://www.dsp.ece.rice.edu/cs/>, 2008.
  - [27] D. Reddy, A. Sankaranarayanan, V. Cevher, and R. Chellappa, “Compressed sensing for multi-view tracking and 3-d voxel reconstruction,” in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2008.
  - [28] E.J.Candes and M.B.Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, pp. 21–30, March 2008.
  - [29] E.J.Candes, “Compressive sampling,” in *Proc. of the International Congress of Mathematicians*, 2006.
  - [30] R.Baraniuk, “Compressive sensing,” *IEEE signal processing magazine*, pp. 118–124, 2007.
  - [31] A.Hormati, O.Roy, Y.M.Lu, and M.Vetterli, “Distributed sampling of signals linked by sparse filtering: theory and applications,” *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1095–1109, March 2010.
  - [32] J. Tropp and A. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, December 2007.
  - [33] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
  - [34] N.Efford, *Digital Image Processing: A Practical Introduction Using JavaTM*. Pearson Education, 2000, ch. Morphological image processing.
  - [35] S.B.Jebara and H.Besbes, “A variable step size filtered sign algorithm for acoustic echo cancellation,” in *IEEE electronic letters*, vol. 39, 2003, pp. 936–93.
  - [36] S.Dhull, S.Arya, and O. Sahu, “Performance variation of lms and its different variants,” *International Journal of Computer Science and Security (IJCSS)*, vol. 4, 2010.
  - [37] H.S.Yazdi, M. Lotfizad, and M.Fathy, “Car tracking by quantised input lms, qx-lms algorithm in traffic scenes,” *Vision, Image and Signal Processing, IEE Proceedings*, vol. 153, no. 1, pp. 37–45, 2006.
  - [38] F. Cheng and Y. Chen, “Real time multiple objects tracking and identification based on discrete wavelet trans-

form,” *Elsevier Pattern Recognition Journal*, vol. 39, p. 1126–1139, 2006.

- [39] “Caviar datasets,” Dataset: EC Funded CAVIAR project/IST 2001 37540, <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>, 2001.
- [40] H. Jiang, W. Deng, and Z. Shen, “Surveillance video processing using compressive sensing,” *arXiv preprint arXiv:1302.1942*, 2013.